

AD 740569

ARL 71-0320  
DECEMBER 1971



## **Aerospace Research Laboratories**

### **A NOTE ON ADF TESTS FOR SUBHYPOTHESES IN MULTIPLE LINEAR REGRESSION**

*M. L. PURI*

*INDIANA UNIVERSITY  
BLOOMINGTON, INDIANA*

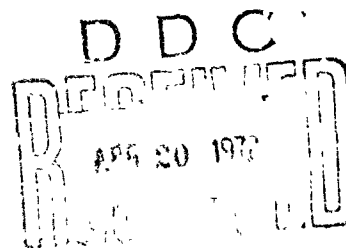
*P. K. SEN*

*UNIVERSITY OF NORTH CAROLINA  
CHAPEL HILL, NORTH CAROLINA*

CONTRACT NO. F33615-71-C-1927  
PROJECT NO. 7071

Approved for public release; distribution unlimited.

NATIONAL TECHNICAL  
INFORMATION SERVICE



AIR FORCE SYSTEMS COMMAND  
**United States Air Force**

## NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Agencies of the Department of Defense, qualified contractors and other government agencies may obtain copies from the

Defense Documentation Center  
Cameron Station  
Alexandria, Virginia 22314

This document has been released to the

CLEARINGHOUSE  
U. S. Department of Commerce  
Springfield, Virginia 22151

for the sale to the public.

Copies of ARL Technical Documentary Reports should not be returned to Aerospace Research Laboratories unless return is required by security considerations, contractual obligations or notices on a specified document.

AIR FORCE: 8-3-72/E00

ACCESSION for	
WSTA	WHITE SECTION <input checked="" type="checkbox"/>
DDC	DIFF SECTION <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
DISTRIBUTION/AVAILABILITY CODES	
DIST	ACRIL
A	

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D		
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)		
1. ORIGINATING ACTIVITY (Corporate author) University of North Carolina Department of Biostatistics Chapel Hill, North Carolina 27514		2a. REPORT SECURITY CLASSIFICATION Unclassified
		2b. GROUP M-2
3. REPORT TITLE  A Note on ADF Tests for Subhypotheses in Multiple Linear Regression		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Scientific Interim		
5. AUTHOR(S) (First name, middle initial, last name)  M. L. Puri and P. K. Sen		
6. REPORT DATE December 1971	7a. TOTAL NO. OF PAGES 8	7b. NO. OF REFS 3
8a. CONTRACT OR GRANT NO. F 33615-71-C-1927 b. PROJECT NO. 7071-00-12 c. DoD Element 61102F d. DoD Subelement 681304		9a. ORIGINATOR'S REPORT NUMBER(S)  9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ARL 71-0320
10. DISTRIBUTION STATEMENT  Approved for public release; distribution unlimited		
11. SUPPLEMENTARY NOTES  TECH OTHER		12. SPONSORING MILITARY ACTIVITY  Aerospace Research Laboratories (LB) Wright-Patterson AFB, Ohio 45433
13. ABSTRACT  The purpose of this note is to show that asymptotically distribution-free (ADF) rank tests proposed by Koul [ <u>Ann. Math. Statist.</u> 41(1970), 1273-1281] for testing subhypotheses in multiple linear regression when based on symmetric scores are asymptotically totally inefficient and can even be inconsistent for symmetric parent distributions. Also, the condition given in his Theorem 2.4 for a test to be ADF is not necessary.		

DD FORM 1 NOV 55 1473

Unclassified

Security Classification

Unclassified  
Security Classification

14	KEY WORDS	LINK A		LINK B		LINK C	
		ROLE	WT	ROLE	WT	ROLE	WT
	Multiple Regression Distribution Free Tests Asymptotic Tests						

Unclassified

Security Classification

# **A NOTE ON ADF TESTS FOR SUBHYPOTHESES IN MULTIPLE LINEAR REGRESSION**

*M. L. PURI*

*INDIANA UNIVERSITY*

*AND*

*P. K. SEN*

*UNIVERSITY OF NORTH CAROLINA*

DECEMBER 1971

CONTRACT F 33615-71-C-1927

PROJECT 7071

Approved for public release; distribution unlimited.

AEROSPACE RESEARCH LABORATORIES  
AIR FORCE SYSTEMS COMMAND  
UNITED STATES AIR FORCE  
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

## FOREWORD

This is an interim report of the work done under Contract F33615-71-C-1927 by the University of North Carolina. The work done by P. K. Sen in this report is sponsored by the Aerospace Research Laboratories under the above contract; it is accomplished on Project 7071 "Research in Applied Mathematics" and is technically monitored by P. R. Krishnaiah. The work of M. L. Puri is supported by the National Science Foundation under Grant NSF-GP-27715.

## ABSTRACT

The purpose of this note is to show that asymptotically distribution-free (ADF) rank tests proposed by Koul [Ann. Math. Statist. 41(1970), 1273-1281] for testing subhypotheses in multiple linear regression when based on symmetric scores are asymptotically totally inefficient and can even be inconsistent for symmetric parent distributions. Also, the condition given in his Theorem 2.4 for a test to be ADF is not necessary.

## TABLE OF CONTENTS

Section	Page
1. Introduction	1
2. Outline of the Proofs	2
References	4



1. Introduction. For every  $n (\geq 1)$ , let  $\{Y_{n1}, \dots, Y_{nn}\}$  be a sequence of independent random variables, such that

$$(1.1) \quad F_{ni}(x) = P\{Y_{ni} \leq x\} = F(x - \beta_1 x_{ni}(1) - \beta_2 x_{ni}(2)), \quad i = 1, \dots, n;$$

$x_{ni}(j)$ ,  $i = 1, \dots, n$ ,  $j = 1, 2$  are known constants, and  $\beta_1, \beta_2$  are unknown parameters. Our problem is to test

$$(1.2) \quad H_0: \beta_1 = 0 \text{ vs. } H_1: \beta_1 \neq 0 \text{ (or } < \text{ or } > 0),$$

where  $\beta_2$  is treated as a nuisance parameter.

Koul(1970) considered an arbitrary estimator  $\hat{\beta}_n$  of  $\beta_2$ , such that  $|n^{\frac{1}{2}}(\hat{\beta}_n - \beta_2)| = O_p(1)$ , and considered the following procedure. Let

$$(1.3) \quad Y_{ni}(b) = Y_{ni} - b x_{ni}(2), \quad \hat{Y}_{ni} = Y_{ni}(\hat{\beta}_n), \quad i = 1, \dots, n,$$

and let  $R_{ni}(b)$  (or  $\hat{R}_{ni}$ ) be the rank of  $Y_{ni}(b)$  (or  $\hat{Y}_{ni}$ ) among the  $n$  values in

$$(1.3). \text{ Let } \bar{x}_n(j) = n^{-1} \sum_{i=1}^n x_{ni}(j), \quad j = 1, 2, \text{ and let}$$

$$(1.4) \quad S_n^{(1)}(b) = n^{-1} \sum_{i=1}^n (x_{ni}(1) - \bar{x}_n(1)) \phi(R_{ni}(b)/(n+1)), \quad \hat{S}_n^{(1)} = S_n^{(1)}(\hat{\beta}_n),$$

where  $\phi(u)$  is absolutely continuous and bounded inside  $[0,1]$ . He showed that under certain conditions,  $\hat{S}_n^{(1)}$  is ADF under  $H_0$  in (1.2), and hence, a test for  $H_0$  can be based on  $\hat{S}_n^{(1)}$ . He showed that if  $F$  is double exponential and  $\phi(u)$  is Freund-Ansary or Mood type (which are both symmetric), then the corresponding  $\hat{S}_n^{(1)}$  is ADF. To this result, we add the following.

Theorem 1. For the entire class of bounded, absolutely continuous and symmetric scores (i.e., for  $\phi(u) = \phi(1-u)$ ,  $u \in [0,1]$ ), Koul's test based on  $S_n^{(1)}$  is ADF, but is totally asymptotically inefficient (in the sense that for local alternatives, the power is equal to the level of significance) when the underlying  $F$  is symmetric.

Theorem 2. Whenever  $\phi(u) = \phi(1-u)$ ,  $u \in [0,1]$ , Koul's ADF tests can also be inconsistent for a fixed  $\beta_1 \neq 0$ , no matter whether  $F$  is symmetric or not.

Let us now define

$$(1.5) \quad S_n^{(2)}(b) = n^{-1} \sum_{i=1}^n (x_{ni}(2) - \bar{x}_n(2)) \phi(R_{ni}(b)/(n+1)),$$

so that when  $\phi(u)$  is non-decreasing,  $S_n^{(2)}(b)$  is non-increasing in  $b$ :  $-\infty < b < \infty$ .

Let then

$$(1.6) \quad \hat{\beta}_{n,1} = \sup\{b: S_n^{(2)}(b) > 0\}, \quad \hat{\beta}_{n,2} = \inf\{b: S_n^{(2)}(b) < 0\};$$

$$(1.7) \quad \hat{\beta}_n = [\hat{\beta}_{n,1} + \hat{\beta}_{n,2}]/2.$$

Finally, let

$$(1.8) \quad \phi(u, f) = -f'(F^{-1}(u))/f(F^{-1}(u)), \quad u \in [0, 1],$$

$$(1.9) \quad b_n(\phi, f) = [n^{-1} \sum_{i=1}^n (x_{ni}(1) - \bar{x}_n(1))(x_{ni}(2) - \bar{x}_n(2))] b(\phi, f);$$

$$(1.10) \quad b(\phi, f) = \int_0^1 \phi(u, f) \phi(u) du = - \int_{-\infty}^{\infty} f'(x) \phi(F(x)) dx.$$

In his theorem 2.4, Koul (1970) has shown that a sufficient condition for  $\hat{S}_n^{(1)}$  to be ADF is that  $b_n(\phi, f) = 0$ . In this context, we have the following:

**Theorem 3.** Whenever  $\phi(u)$  is monotonic in  $u \in [0, 1]$ , and  $\hat{\beta}_n$  is defined by (1.6)-(1.7),  $\hat{S}_n^{(1)}$  is ADF, no matter whether  $b_n(\phi, f)$  is equal to 0 or not.

The proofs of the theorems are sketched in section 2. It clearly follows from the first two theorems that symmetric scores (such as Mood's where  $\phi(u) = (u - \frac{1}{2})^2$ ) are unsuitable for this problem, while from the last theorem, it follows that for monotonic scores (such as Wilcoxon's where  $\phi(u) = u - 1/2$ ), Koul's condition that  $b_n(\phi, f) = 0$  can easily be avoided when the estimator  $\hat{\beta}_n$  is based on the same scores.

2. Outline of the proofs. (i) Theorem 1.  $\phi$  symmetric implies that  $\phi(u) = \phi(1-u)$ ,  $u \in [0, 1]$ , while  $F$  symmetric implies that  $\phi(u, f) = -\phi(1-u, f)$ ,  $u \in [0, 1]$ . Hence, by (1.10),

$$(2.1) \quad b(\phi, f) = 0 \quad \Rightarrow \quad b_n(\phi, f) = 0.$$

Consider then the sequence of local alternatives  $\{H_n\}$ , where

$$(2.2) \quad H_n: \beta_1 = \beta_1^{(n)} = n^{-1/2}\theta, \quad \theta \neq 0.$$

Now, by Theorem 2.3 of Koul (1970), under (2.1),

$$(2.3) \quad n^{1/2} \hat{S}_n^{(1)} = n^{1/2} S_n^{(1)}(\beta_2) + o_p(1), \text{ as } n \rightarrow \infty,$$

where  $S_n^{(1)}(\beta_2)$  is defined by (1.4) for  $b = \beta_2$ . Since by (1.1) and (1.3),  $Y_{ni}(\beta_2)$  has the distribution  $F(x - \beta_1 x_{ni}(1))$ , for  $i = 1, \dots, n$ , by the well-known results of Hájek(1962), under (2.2),

$$(2.4) \quad \mathcal{L}(n^{1/2} S_n^{(1)}(\beta_2) - \theta b_n(\phi, f) \mid H_n) \rightarrow \mathcal{N}(0, A^2),$$

where

$$(2.5) \quad A^2 = \int_0^1 \phi^2(u) du - \left( \int_0^1 \phi(u) du \right)^2.$$

As shown in (2.1), under the hypothesis of theorem 1,  $b_n(\phi, f) = 0$ , and hence, by (2.3) and (2.4), we conclude that for all finite  $\theta$ , the asymptotic power of the test based on  $\hat{S}_n^{(1)}$  remains equal to  $\alpha$ , the level of significance. That is, the test is ADF but totally inefficient. Q.E.D.

(ii) Theorem 2. Let  $\tau_\alpha$  be the upper 100 $\alpha$ % point of the standard normal distribution.

Then, we are to show that

$$(2.6) \quad P\{ |n^{1/2} \hat{S}_n^{(1)}| > A\tau_{\alpha/2} \mid \beta_1 \neq 0 \}$$

does not necessarily converge to 1 as  $n \rightarrow \infty$ . This we show by means of the following example.

Let  $n = 2m$ ,  $x_{n1}(1) = \dots = x_{nm}(1) = -1$ ,  $x_{nm+1}(1) = \dots = x_{nn}(1) = 1$ ,  $\beta_1 = 2 + \epsilon$ ,  $\epsilon > 0$ , and let  $F(x)$  be such that  $F(-1) = 0$ ,  $F(1) = 1$ , [the later can always be made by choosing first an absolutely continuous df  $G$  with finite Fisher information, and letting then  $F(x) = [G(x) - G(-1)]/[G(1) - G(-1)]$  for  $x \in [-1, 1]$ , and  $F(x) = 0$  for  $x \leq -1$ ,  $F(x) = 1$ , for  $x \geq 1$ .] Then, proceeding as in Koul(1970), we have

$$(2.7) \quad \max_{1 \leq i \leq n} |Y_{ni} - \hat{\beta}_n x_{ni}(2)| = o_p(1), \text{ as } n \rightarrow \infty,$$

and hence, for every  $\epsilon > 0$ , there exists an  $n_\epsilon$ , such that for  $n \geq n_\epsilon$ , the right hand side of (2.7) is less than  $\epsilon$ , with probability approaching to unity, as  $n \rightarrow \infty$ .

Since the variation of  $F$  is contained entirely within the range  $(-1, 1)$  and  $\beta_1 = 2 + \epsilon$ ,  $\epsilon > 0$ , it follows that for  $n \geq n_\epsilon$ ,

$$(2.8) \quad Y_{ni} - \hat{\beta}_n x_{ni}(2) < Y_{nj} - \hat{\beta}_n x_{nj}(2), \quad i=1, \dots, m, \quad j=m+1, \dots, n,$$

with a probability arbitrarily close to unity as  $n \rightarrow \infty$ . Hence,

$$(2.9) \quad P\left\{ \left| \hat{S}_n^{(1)} - n^{-1} \sum_{i=1}^m [\phi((i+m)/2m) - \phi(i/2m)] \right| = 0 \right\} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

Since  $\phi(u) = \phi(1-u)$  for all  $u \in [0,1]$ , by (2.9),

$$(2.10) \quad n^{1/2} \hat{S}_n^{(1)} = 0 \text{ with probability tending to 1 as } n \rightarrow \infty,$$

which shows that (2.6) does not converge to 1 as  $n \rightarrow \infty$ , so the test is inconsistent.

The same proof holds for  $\beta_1 < -2$  or for  $|\beta_1| > 2$ . Q.E.D.

(iii) Theorem 3. Let  $c_{jj'}^{(n)} = n^{-1} \sum_{i=1}^n (x_{ni}(j) - \bar{x}_n(j))(x_{ni}(j') - \bar{x}_n(j'))$ , for  $j, j' = 1, 2$ . Then, by Theorem 2.3 of Koul (1970), as  $n \rightarrow \infty$ ,

$$(2.11) \quad n^{1/2} [\hat{S}_n^{(1)} - s_n^{(1)}(\beta_2)] = -n^{1/2} (\hat{\beta}_n - \beta_2) b_n(\phi, f) + o_p(1).$$

On the other hand, under (1.2),  $\beta_1 = 0$ , so that when  $\hat{\beta}_n$  is defined by (1.6) and (1.7), by the fundamental results of Jurečková (1969), as  $n \rightarrow \infty$ ,

$$(2.12) \quad n^{1/2} | s_n^{(2)}(\beta_2) - (\hat{\beta}_n - \beta_2) b_n(\phi, f) | = o_p(1).$$

Consequently, by (2.11) and (2.12), under (1.2), as  $n \rightarrow \infty$ ,

$$(2.13) \quad n^{1/2} \hat{S}_n^{(1)} = n^{1/2} [s_n^{(1)}(\beta_2) - s_n^{(2)}(\beta_2)] + o_p(1).$$

Finally, under (1.2),  $n^{1/2} [s_n^{(1)}(\beta_2), s_n^{(2)}(\beta_2)]$  jointly has asymptotically a bivariate normal distribution with mean  $(0,0)$  and covariance matrix  $A^2 \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ , where  $c_{jj'} = \lim_{n \rightarrow \infty} c_{jj'}^{(n)}$ ,  $j, j' = 1, 2$ . Hence, by (2.13), under  $H_0$  in (1.2),

$$(2.14) \quad \mathcal{L}_n(n^{1/2} \hat{S}_n^{(1)}) \rightarrow \mathcal{N}(0, A^2(c_{11} + c_{22} - 2c_{12})), \text{ as } n \rightarrow \infty,$$

and as  $A^2$  as well as the  $c_{jj'}$  are known quantities, the result follows. Q.E.D.

#### REFERENCES

- [1] HÁJEK, J. (1962). Asymptotically most powerful rank order tests. Ann. Math. Statist. 33, 1124-1147.
- [2] JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. Ann. Math. Statist. 40, 1889-1900.
- [3] KOUL, H.L. (1970). A class of ADF tests for subhypotheses in multiple linear regression. Ann. Math. Statist. 41, 1273-1281.